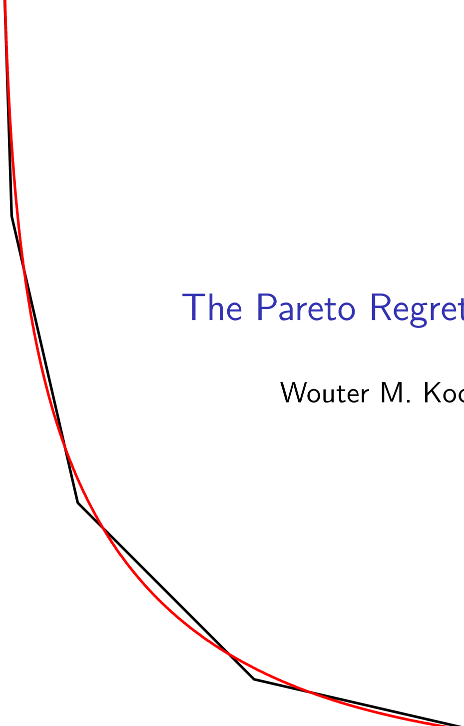
The Pareto Regret Frontier

Wouter M. Koolen



The online learning philosophy

Want to solve a learning problem?

The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches ☺

The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches ☺
  - models from literature ☹

The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches ☺
  - models from literature ☹
  - students ($)

The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches ☺
  - models from literature ☹
  - students ($)
  - employees ($$)

The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches ☺
  - models from literature ☹
  - students ($)
  - employees ($$)
  - companies ($$$)

The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches ☺
  - models from literature ☹
  - students ($)
  - employees ($$)
  - companies ($$$)
  - ...

The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches ☺
  - models from literature ☹
  - students ($)
  - employees ($$)
  - companies ($$$)
  - ...
- Each "expert" implements a solution

The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches ☺
  - models from literature ☹
  - students (\$)
  - employees (\$\$)
  - companies (\$\$\$)
  - ...
- Each "expert" implements a solution
- Use aggregation algorithm to combine solutions in production

The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches ☺
  - models from literature ☹
  - students ($)
  - employees ($$)
  - companies ($$$)
  - ...
- Each “expert” implements a solution
- Use aggregation algorithm to combine solutions in production
- Profit!

That aggregation algorithm

T rounds, K experts

$$\underbrace{L_T - \min_k L_T^k}_{\text{regret}} \leq \sqrt{T/2 \ln K}$$

That aggregation algorithm

T rounds, K experts

$$\underbrace{L_T - \min_k L_T^k}_{\text{regret}} \leq \sqrt{T/2 \ln K}$$

What if

- Lots of experts? shotgun-style "throw in all you got"
- Special experts? company's current strategy

Regret as a multi-objective criterion

A vector $\langle r_1 \dots r_K \rangle$ is T-**realisable** if there is a strategy ensuring

$$L_T - L_T^k \;\leq\; r_k \qquad \text{for each expert } k$$

Regret as a multi-objective criterion

A vector $\langle r_1 \ldots r_K \rangle$ is T-**realisable** if there is a strategy ensuring

$$L_T - L_T^k \;\leq\; r_k \qquad \text{for each expert } k$$

Suggestive: for every expert prior $\mathbb{P}$, the following is T-realisable:

$$\left\langle \sqrt{T/2\,(-\ln \mathbb{P}(k))} \right\rangle_{k=1}^{K} \qquad ????$$

Regret as a multi-objective criterion

A vector $\langle r_1 \ldots r_K \rangle$ is T-**realisable** if there is a strategy ensuring

$$L_T - L_T^k \;\leq\; r_k \qquad \text{for each expert } k$$

Suggestive: for every expert prior $\mathbb{P}$, the following is T-realisable:

$$\left\langle \sqrt{T/2\,(-\ln \mathbb{P}(k))} \right\rangle_{k=1}^{K} \qquad ????$$

but this is false

Regret as a multi-objective criterion

A vector $\langle r_1 \ldots r_K \rangle$ is T-**realisable** if there is a strategy ensuring

$$L_T - L_T^k \;\leq\; r_k \qquad \text{for each expert } k$$

Suggestive: for every expert prior $\mathbb{P}$, the following is T-realisable:

$$\left\langle \sqrt{T/2\,(-\ln \mathbb{P}(k))} \right\rangle_{k=1}^{K} \qquad ????$$

but this is false

So what *can* be realised?

Results in a nutshell

- Absolute loss (or $K = 2$ experts)
  - Exact results
    - Characterisation of T-realisable frontier (combinatorial)
    - Strategy for each trade-off

Results in a nutshell

- Absolute loss (or $K = 2$ experts)
  - Exact results
    - Characterisation of T-realisable frontier (combinatorial)
    - Strategy for each trade-off
  - Asymptotic (large T):
    - Smooth limit frontier
    - Smooth limit strategy
    - For any $p \in [0, 1]$ we can realise:

$$\left\langle \sqrt{T \ (-\ln(p))}, \sqrt{T \ (-\ln(1-p))} \right\rangle$$

but we can do better(!)

Results in a nutshell

- Absolute loss (or $K = 2$ experts)
  - Exact results
    - Characterisation of T-realisable frontier (combinatorial)
    - Strategy for each trade-off
  - Asymptotic (large T):
    - Smooth limit frontier
    - Smooth limit strategy
    - For any $p \in [0,1]$ we can realise:

$$\left\langle \sqrt{T\ (-\ln(p))}, \sqrt{T\ (-\ln(1-p))} \right\rangle$$

      but we can do better(!)

- For $K > 2$ experts
  - For every expert prior $\mathbb{P}(k)$, we can realise

$$\left\langle \sqrt{2.6\, T\ (-\ln \mathbb{P}(k))} \right\rangle_{k=1}^{K}$$

    using a recursive combination of 2-expert algorithms.

Absolute loss game

Each round $t \in \{1, \ldots, T\}$ the learner assigns a probability $p_t \in [0, 1]$ to the next outcome being a 1, after which the actual outcome $x_t \in \{0, 1\}$ is revealed, and the learner suffers

$$\textbf{absolute loss} \qquad |p_t - x_t|.$$

The **regret** w.r.t. the strategy that always predicts $k \in \{0, 1\}$ is

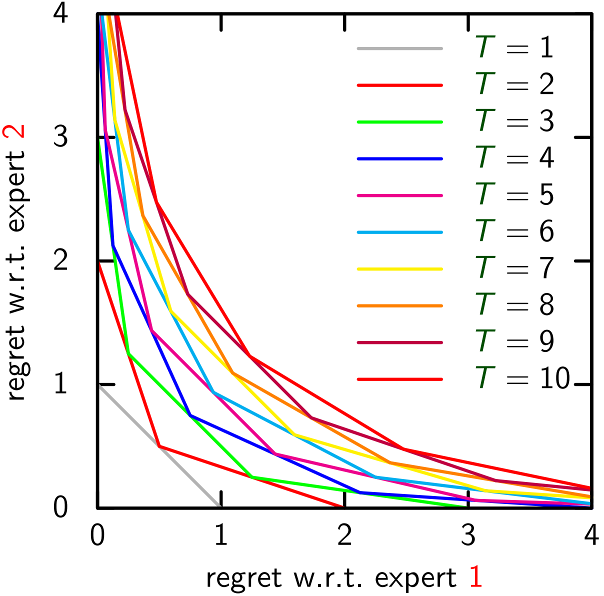
$$R_T^k \ := \ \sum_{t=1}^{T} \big(|p_t - x_t| - |k - x_t|\big)$$

A candidate trade-off $\langle r_0, r_1 \rangle \in \mathbb{R}^2$ is called T-**realisable** for the T-round absolute loss game if there is a strategy that keeps the regret w.r.t. each $k \in \{0, 1\}$ below r_k, i.e. if

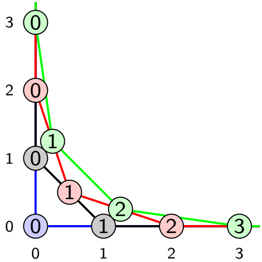
$$\exists p_1 \forall x_1 \cdots \exists p_T \forall x_T : R_T^0 \leq r_0 \text{ and } R_T^1 \ \leq r_1$$

We denote the set of all T-realisable trade-offs by $\mathbb{G}_T$.

The set $\mathbb{G}_T$, i.e. the T-realisable tradeoffs $\langle r_0, r_1 \rangle$



Pareto Frontier and Optimal Strategy



Theorem

The Pareto frontier of $\mathbb{G}_T$ is piece-wise linear with $T+1$ vertices:

$$\left\langle f_T(i), f_T(T-i) \right\rangle \quad 0 \le i \le T \quad \textit{where} \quad f_T(i) := \sum_{j=0}^{i} j 2^{j-T} \binom{T-j-1}{T-i-1}.$$

The optimal strategy at vertex i assigns to $x = 1$ probability

$$p_T(0) := 0, \quad p_T(T) := 1, \qquad p_T(i) \; := \; \frac{f_{T-1}(i) - f_{T-1}(i-1)}{2} \quad 0 < i < T,$$

and it interpolates linearly in between consecutive vertices.

Asymptotic analysis

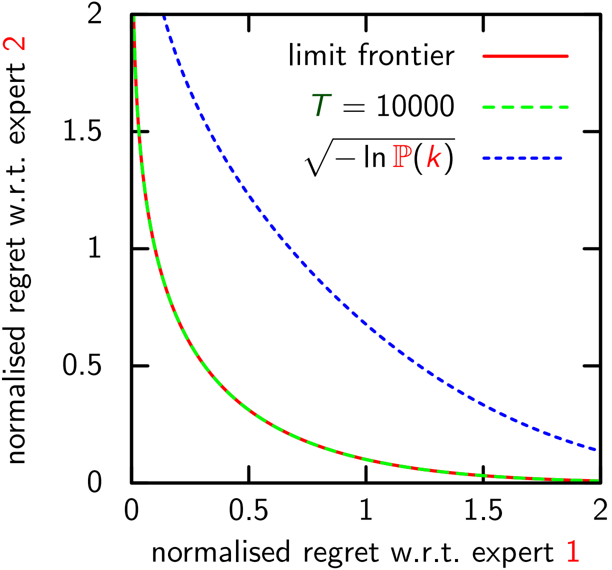
Idea: normalise and then make T large

Asymptotic analysis

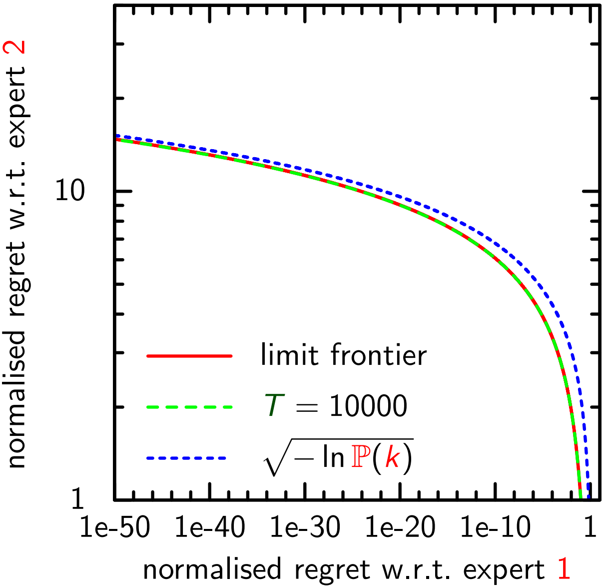
Idea: normalise and then make T large

$$\mathbb{G} \quad := \quad \lim_{T \to \infty} \frac{\mathbb{G}_T}{\sqrt{T}}.$$

Asymptotic plot (moderate)



Asymptotic plot (tail)



Asymptotic Pareto Frontier and Optimal Strategy

Theorem

The Pareto frontier of $\mathbb{G}$ *is the smooth curve*

$$\left\langle f(u), f(-u) \right\rangle \quad u \in \mathbb{R}, \quad \textit{where} \quad f(u) := \int_{-\infty}^{u} \Phi(x)\,\mathrm{d}x,$$

and $\Phi(u) \;:=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{x^2}{2}}\,\mathrm{d}x$ *is the standard normal CDF . The optimal strategy converges to*

$$p(u) \;=\; \Phi(u).$$

Use of asymptotics

- Smooth formula easier to handle
- Allows us to appreciate that sqrt-min-log-prior tradeoffs are realisable with constant 1 (not $1/\sqrt{2}$) ...
- ... but have suboptimal lower-order terms.
- Suggests smoothened algorithms

$K > 2$ experts

Combine algorithm for $K = 2$ into unbalanced binary tree.

Outermost algorithm combines expert with least prior vs rest

Gives us

$$\sqrt{2.6\,T\,(-\ln \mathbb{P}(k))}$$

$K > 2$ experts

Combine algorithm for $K = 2$ into unbalanced binary tree.

Outermost algorithm combines expert with least prior vs rest

Gives us

$$\sqrt{2.6\,T\,(-\ln \mathbb{P}(k))}$$

But we would like to have the exact Pareto frontier.

Conclusion

- We need unfair regret bounds
- Reinterpret regret as a multi-criterion objective
- Exact Pareto frontier for $K = 2$ experts
- with optimal algorithm
- And useful formula for asymptotic Pareto frontier
- with asymptotic algorithm
- Trick for $K > 2$ experts